Table of Contents

Appendix Figure Legends

Appendix Figure S1.
A. Principal component analysis (PCA) of protein-coding gene expression levels, for untreated CHC and control samples. The log2-transformed RPKM levels of all expressed protein-coding genes were used as input for the PCA (Materials and Methods). Black: non-CHC; green: CHC low ISG; red: CHC high ISG samples.
B. Same as **A**, for protein-coding genes associated with "response to interferon" Gene Ontology terms (Materials and Methods).

Appendix Figure S2.
A. Results of the Homer transcription factor binding enrichment analysis (Materials and Methods), for protein-coding genes that are up-regulated in CHC low ISG compared to control non-CHC samples. The numbers in parentheses represent the percentage of genes that have this motif in the promoter region and the estimated enrichment with respect to the genomic background.
B. Boxplots of the intensity of the expression change (log2 fold) between non-CHC patients and CHC low ISG patients, for different categories of protein-coding genes defined based on the miRNAs that are predicted to target them. Target predictions were determined with HITS-CLIP in Huh7 cells (Luna et al, 2015). From left to right: genes that are not targets of any expressed miRNA; genes that are targeted by at least one expressed miRNA; genes that are targeted by miR-122-5p, miR-192/215, let7a-i/miR-98 and miR-15, respectively.
C. Boxplots of the intensity of the expression change (log2 fold) between non-CHC patients and CHC low ISG patients, for different categories of protein-coding genes defined based on the number of miRNA families that are predicted to target them. Only miRNAs with high expression in normal or HCV-infected liver were analyzed (Hou et al, 2011). Evolutionarily conserved miRNA target predictions were extracted from TargetScan v7.1 (Agarwal et al, 2015) (Materials and Methods).

Appendix Figure S3.
A. Results of the Homer transcription factor binding enrichment analysis (Materials and Methods), for protein-coding genes that are up-regulated in CHC high ISG compared to control non-CHC samples. The numbers in parentheses represent the percentage of genes that have this motif in the promoter region and the estimated enrichment with respect to the genomic background.
B. Same as **A**, for genes down-regulated in CHC high ISG compared to control non-CHC samples.
C. Barplots representing the expression patterns of three categories of genes (all expressed genes, genes up-regulated in CHC high ISG patients, and genes down-regulated in CHC high ISG patients compared to control non-CHC samples), in the GTEx tissue transcriptome collection. For each expressed gene, we scored the tissue or cell type in which its maximum expression level was reached. The height of the rectangles represents the percentage of genes that reaches maximum expression in each of the tissues.

Appendix Figure S4.
A. Volcano plot for the differential expression analysis between CHC patients with low levels of endogenous ISG activation (CHC low ISG) and CHC patients with high levels of endogenous ISG activation (CHC high ISG). The X-axis represents the log2 fold expression change in CHC high ISG compared to CHC low patients. The Y-axis represents the false discovery rate (with a -log10 transformation) of the differential expression test. Protein-coding genes, candidate long non-coding RNAs (lncRNAs) and other gene categories (including pseudogenes and transcripts with unclear coding potential, Materials and Methods) are represented separately. Only genes with false discovery rate (FDR) < 10% and with

fold expression change > 1.5 are shown.

B. Heatmap of the expression patterns of the top differentially expressed genes between CHC low ISG and CHC high ISG patients. To select the top differentially expressed genes, we set a minimum absolute fold change threshold of 3.5 for up-regulated genes and of 2 for down-regulated genes. The heatmap represents the Z-score of the log2-transformed RPKM (reads per kilobase of exon per million mapped reads) gene expression levels, normalized based on housekeeping genes (Materials and Methods).

C. Results of the Homer transcription factor binding enrichment analysis (Materials and Methods), for protein-coding genes that are up-regulated in CHC high ISG compared to CHC low ISG samples. The numbers in parentheses represent the percentage of genes that have this motif in the promoter region and the estimated enrichment with respect to the genomic background.

Appendix Figure S5.

A. Results of the Homer transcription factor binding enrichment analysis (Materials and Methods), for protein-coding genes that are up-regulated following 4h of pegIFNa/ribavirin treatment. The numbers in parentheses represent the percentage of genes that have this motif in the promoter region and the estimated enrichment with respect to the genomic background.

B. Same as **A**, for genes that are up-regulated following 16h of pegIFNa/ribavirin treatment.

C. Same as **A**, for genes that are up-regulated following 48h of pegIFNa/ribavirin treatment.

D. Same as **A**, for genes that are up-regulated following 96h of pegIFNa/ribavirin treatment.

E. Same as **A**, for genes that are up-regulated following 144h of pegIFNa/ribavirin treatment.

F. RNA-seq profiles along the NRIR lncRNA, for pre-treatment and post-treatment biopsies at the 4h time point. Gray: normalized read coverage for baseline/pre-treatment biopsies. Black: normalized read coverage for post-treatment biopsies. Red: positive difference between post-treatment and pre-treatment biopsies. Green: negative difference between post-treatment and pre-treatment biopsies.

G. Same as **F**, for the lncRNA CTD-2521M24.5.

H. Same as **F**, for the lncRNA CTD-2095E4.5.

Appendix Figure S6.

A. Differential expression patterns of the genes that are significantly up-regulated (FDR < 0.1 and minimum absolute fold change 1.5) in CHC high ISG patients compared to CHC low ISG patients. Blue dots: genes also significantly differentially expressed following pegIFNa/ribavirin treatment (FDR <0.05); red dots: genes significant only in the comparison between low ISG and high ISG samples; black dots: genes not tested in the pegIFNa/ribavirin comparison due to low or highly variable expression levels (Materials and Methods). X-axis: log2 fold expression change in pegIFNa-treated compared to control biopsies. Y-axis: log2 fold expression change in high ISG patients compared to low ISG patients. The numbers of the genes in each category are depicted on the plot area, with the same color code.

B. Similar to **A**, for genes that are significantly down-regulated (FDR < 0.1 and minimum absolute fold change 1.5) in CHC low ISG and CHC high ISG patients.

C. Histogram of the difference in expression levels (log2-transformed RPKM) between samples treated with pegIFNa/ribavirin and CHC high ISG samples, for the genes that are up-regulated (FDR < 0.1 and minimum absolute fold change 1.5) in CHC high ISG patients compared to CHC low ISG patients. The differences were computed between expression levels averaged across all relevant samples.

D. Similar to **C**, for genes that are significantly down-regulated (FDR < 0.1 and minimum absolute fold change 1.5) in CHC high ISG patients compared to CHC low ISG patients.

Appendix Figure S7.

A. Principal component analysis (PCA) of protein-coding gene expression levels, for treated and untreated CHC. The log2-transformed RPKM levels of protein-coding genes that are differentially expressed following pegIFNa treatment (minimum absolute fold change 2, FDR<0.05, at any of the 5 time points) were used as input for the PCA. Green: CHC low ISG, before treatment; orange: CHC high ISG samples, before treatment; red: pegIFNa/ribavirin-treated samples. The dot shapes depict the time points of the pegIFNa treatment.

B. Hierarchical clustering analysis of treated and untreated CHC samples. For the hierarchical clustering, pairwise distances between samples were computed as 1-Pearson's correlation coefficient, using protein-coding genes that are differentially expressed following pegIFNa treatment (minimum absolute fold change 2, FDR<0.05, at any of the 5 time points). The heatmap represents pairwise distances between samples. Cold colors (blue) represent small distances; warm colors (yellow) represent large distances.

Appendix Figure S8.

Heatmap of the expression patterns of 100 selected protein-coding genes, in control non-CHC samples, CHC low ISG and high ISG samples (pre-treatment) and pegIFNa/ribavirin-treated samples. The selected genes were strongly differentially expressed (minimum absolute fold change 2, FDR<0.01) following pegIFNa/ribavirin treatment for at least 2 time points, but were not significantly differentially expressed between low ISG and high ISG samples (FDR >= 0.1). With these settings, 85 up-regulated and 25 down-regulated genes were included. The heatmap represents the Z-score of the log2-transformed RPKM expression levels. Cold colors represent low expression values; warm colors represent high expression values.

Appendix Figure S9.

A. RNA-seq profiles along the miR-331/miR-3685 host gene, for pre-treatment and post-treatment biopsies at the 16h time point, for the three analyzed individuals (I1, I2 and I3). Gray: normalized read coverage for baseline/pre-treatment biopsies. Black: normalized read coverage for post-treatment biopsies. Red: positive difference between post-treatment and pre-treatment biopsies. Green: negative difference between post-treatment and pre-treatment biopsies.

B. Same as **A**, for the miR-192/miR-194 host gene.

C. Same as **A**, for the miR-146a host gene.

Appendix Figure S10.

Quantification of mature miRNA levels by qPCR (Materials and methods). We note that qPCR could not be performed for some samples due to low remaining RNA quantity (Dataset EV 12).

A. Mean Ct value (average of 3 technical replicates, normalized by total RNA quantity) for miR-122-5p, before and after pegIFNa/ribavirin treatment, at the 5 time points. Black dots: before treatment; red dots: after treatment. Data from paired biopsies are connected with a dotted line.

B. Log2 fold change (difference between Ct values before and after treatment) for miR-122-5p. The dotted red line represents a null fold change. Each dot represents a pair of pre-treatment and post-treatment biopsies.

C. Same as **A**, for miR-146a-5p.

D. Same as **B**, for miR-146a-5p.

E. Same as **A**, for miR-331-3p.

F. Same as **B**, for miR-331-3p.

Appendix Figure S11.
A. Dot-chart of the median fold expression change following pegIFNa/ribavirin treatment at the 4h time point, for predicted targets of miRNAs whose hosts are down-regulated (red dots) or of other miRNAs (black dots). We show 79 miRNA families with at least 100 conserved target genes. As a control, we show the median fold expression change for all genes predicted to be targets of any liver-expressed miRNAs or of genes not predicted to be targeted by these miRNAs (blue dots). The horizontal bars represent 95% confidence intervals of the median, constructed with a bootstrap resampling approach (Materials and Methods). Only miRNAs with high expression in normal or HCV-infected liver were analyzed (Hou et al, 2011). Evolutionarily conserved miRNA target predictions were extracted from TargetScan v7.1 (Agarwal et al, 2015) (Materials and Methods). miRNAs highlighted in red correspond to primary miRNA transcripts significantly down-regulated at the 4h time point; miRNAs highlighted in orange correspond to primary miRNA transcripts significantly down-regulated at other time points.
B. Same as **A**, for the 48h time point.

Appendix Figure S12.
A. Dot-chart of the median fold expression change following pegIFNa/ribavirin treatment at the 96h time point, for predicted targets of miRNAs whose hosts are down-regulated (red dots) or of other miR-NAs (black dots). We show 79 miRNA families with at least 100 conserved target genes. As a control, we show the median fold expression change for all genes predicted to be targets of any liver-expressed miRNAs or of genes not predicted to be targeted by these miRNAs (blue dots). The horizontal bars represent 95% confidence intervals of the median, constructed with a bootstrap resampling approach (Materials and Methods). Only miRNAs with high expression in normal or HCV-infected liver were analyzed (Hou et al, 2011). Evolutionarily conserved miRNA target predictions were extracted from TargetScan v7.1 (Agarwal et al, 2015) (Materials and Methods). miRNAs highlighted in red correspond to primary miRNA transcripts significantly down-regulated at the 96h time point; miRNAs highlighted in orange correspond to primary miRNA transcripts significantly down-regulated at other time points.
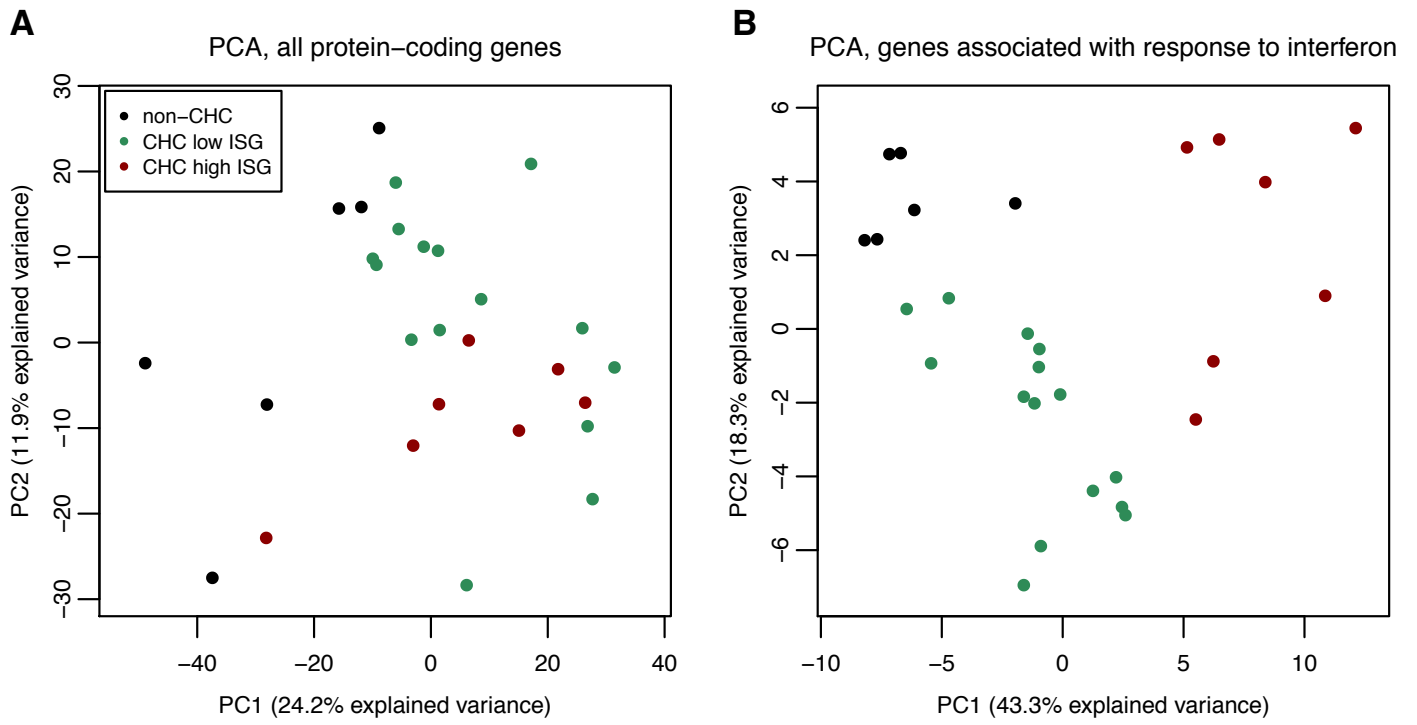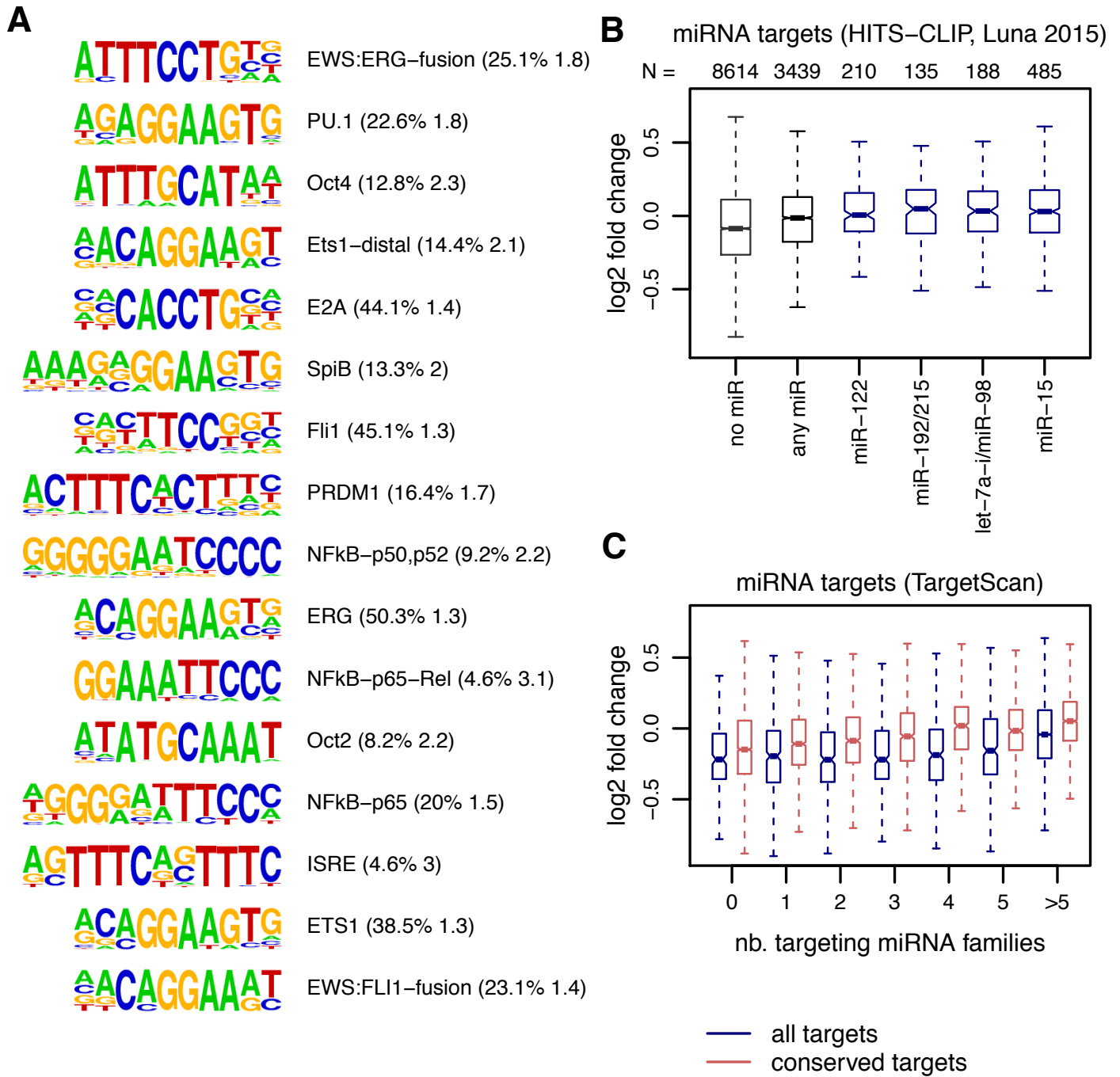**B.** Same as **A**, for the 144h time point.

Appendix Figure S1

A

PCA, all protein-coding genes

B

PCA, genes associated with response to interferon

A

EWS:ERG−fusion (25.1% 1.8)

PU.1 (22.6% 1.8)

Oct4 (12.8% 2.3)

Ets1−distal (14.4% 2.1)

E2A (44.1% 1.4)

SpiB (13.3% 2)

Fli1 (45.1% 1.3)

PRDM1 (16.4% 1.7)

NFkB−p50,p52 (9.2% 2.2)

ERG (50.3% 1.3)

NFkB−p65−Rel (4.6% 3.1)

Oct2 (8.2% 2.2)

NFkB−p65 (20% 1.5)

ISRE (4.6% 3)

ETS1 (38.5% 1.3)

EWS:FLI1−fusion (23.1% 1.4)

B

C

A

AGTTTCAGTTTC ISRE (13.6% 9.6)

GAAAGTGAAAGT IRF1 (17.4% 5.9)

GAAAGTGAAAGT IRF2 (14.9% 6.1)

ACTTTCGTTTCT T1ISRE (2.6% 17.1)

CGGAAGTGAAAC PU.1−IRF (40.3% 1.5)

ACTTTCACTTTC PRDM1 (18.2% 1.9)

ACTGAAACCA IRF4 (14.4% 2.1)

CAGTTTCATTATGACTCA bZIP:IRF (12.1% 2)

AGGGGATTTCCC NFkB−p65 (17.7% 1.6)

AACAGGAAGT Ets1−distal (10.8% 1.8)

GGAAATTCCC NFkB−p65−Rel (3.8% 2.8)

ATTTCCTGTG EWS:ERG−fusion (18.7% 1.4)

B

CAGACGTCAAAGTCCA HNF4a (19.4% 2.3)

GGTTAAACATTAA Hnf1 (6.7% 4.2)

C

Expression pattern (GTex)

tissue w. max expression (% genes)

N = 11806 319 171
all up down

- liver
- spleen
- lymphocytes
- whole blood
- brain
- adipose
- other
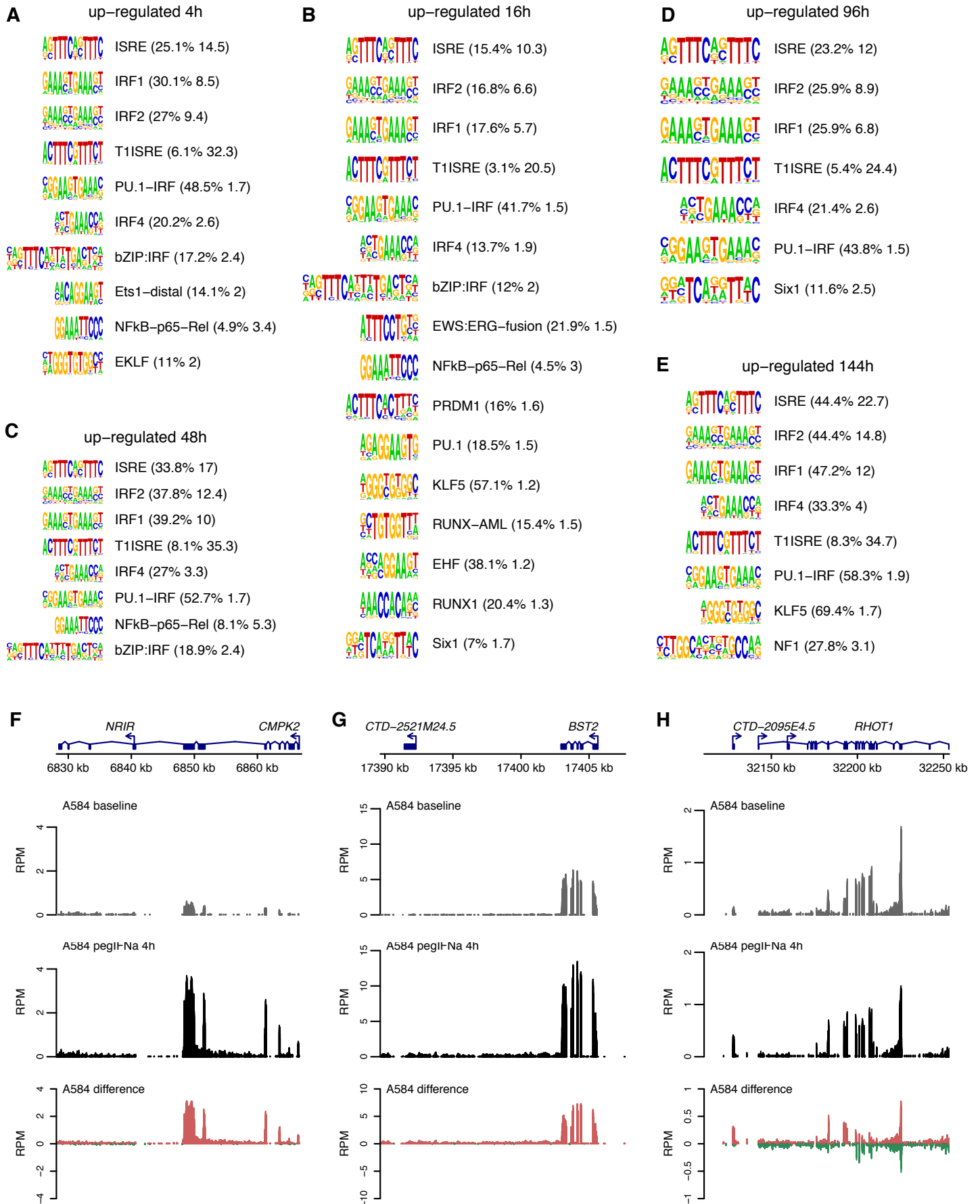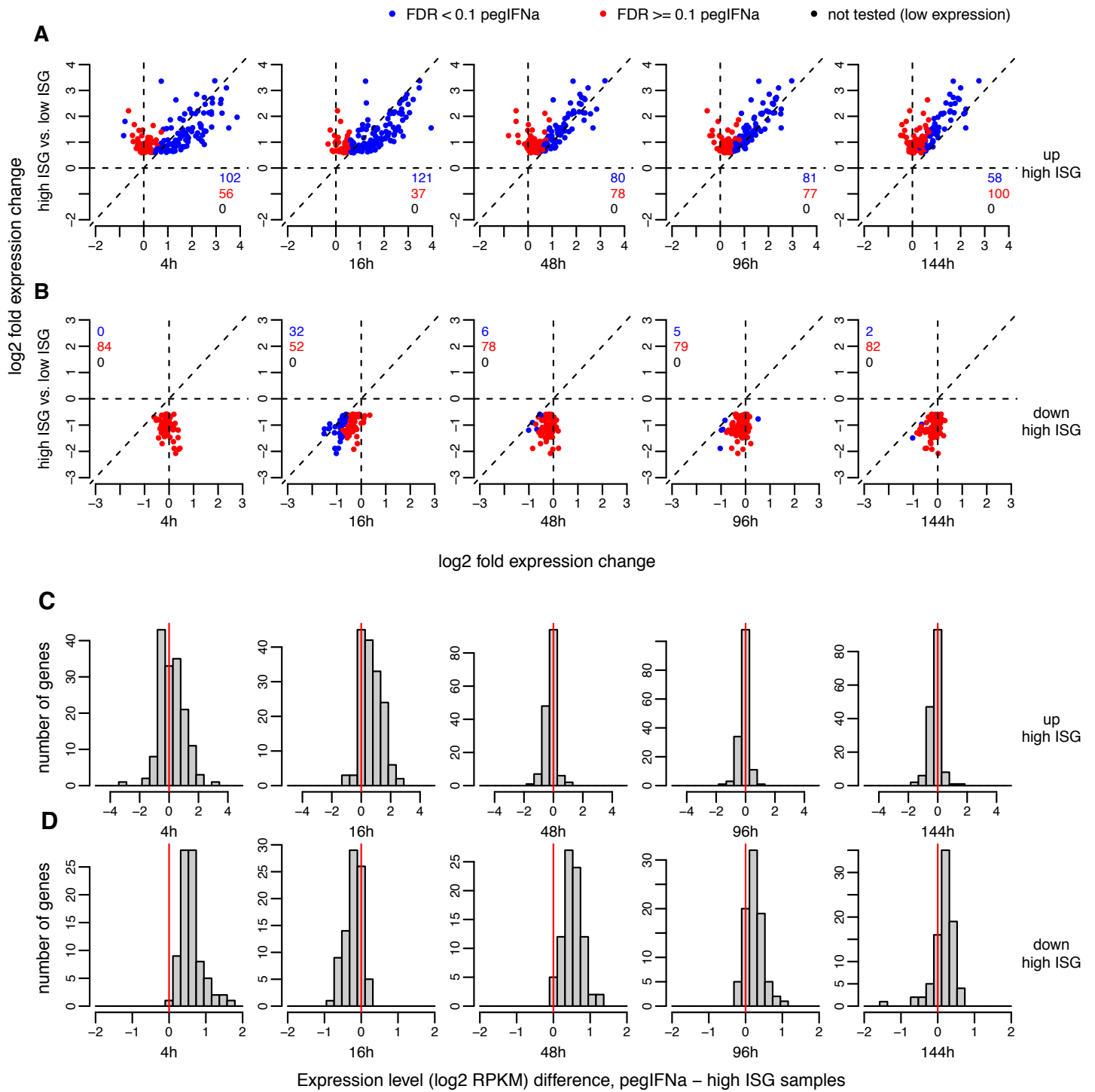
Differential expression, CHC low ISG vs. CHC high ISG patients

Appendix Figure S5

A up–regulated 4h

ISRE (25.1% 14.5)
IRF1 (30.1% 8.5)
IRF2 (27% 9.4)
T1ISRE (6.1% 32.3)
PU.1–IRF (48.5% 1.7)
IRF4 (20.2% 2.6)
bZIP:IRF (17.2% 2.4)
Ets1–distal (14.1% 2)
NFkB–p65–Rel (4.9% 3.4)
EKLF (11% 2)

C up–regulated 48h

ISRE (33.8% 17)
IRF2 (37.8% 12.4)
IRF1 (39.2% 10)
T1ISRE (8.1% 35.3)
IRF4 (27% 3.3)
PU.1–IRF (52.7% 1.7)
NFkB–p65–Rel (8.1% 5.3)
bZIP:IRF (18.9% 2.4)

B up–regulated 16h

ISRE (15.4% 10.3)
IRF2 (16.8% 6.6)
IRF1 (17.6% 5.7)
T1ISRE (3.1% 20.5)
PU.1–IRF (41.7% 1.5)
IRF4 (13.7% 1.9)
bZIP:IRF (12% 2)
EWS:ERG–fusion (21.9% 1.5)
NFkB–p65–Rel (4.5% 3)
PRDM1 (16% 1.6)
PU.1 (18.5% 1.5)
KLF5 (57.1% 1.2)
RUNX–AML (15.4% 1.5)
EHF (38.1% 1.2)
RUNX1 (20.4% 1.3)
Six1 (7% 1.7)

D up–regulated 96h

ISRE (23.2% 12)
IRF2 (25.9% 8.9)
IRF1 (25.9% 6.8)
T1ISRE (5.4% 24.4)
IRF4 (21.4% 2.6)
PU.1–IRF (43.8% 1.5)
Six1 (11.6% 2.5)

E up–regulated 144h

ISRE (44.4% 22.7)
IRF2 (44.4% 14.8)
IRF1 (47.2% 12)
IRF4 (33.3% 4)
T1ISRE (8.3% 34.7)
PU.1–IRF (58.3% 1.9)
KLF5 (69.4% 1.7)
NF1 (27.8% 3.1)

F NRIR   CMPK2

6830 kb 6840 kb 6850 kb 6860 kb

A584 baseline
A584 pegIF-Na 4h
A584 difference

G CTD–2521M24.5   BST2

17390 kb 17395 kb 17400 kb 17405 kb

A584 baseline
A584 pegIF-Na 4h
A584 difference

H CTD–2095E4.5   RHOT1

32150 kb 32200 kb 32250 kb

A584 baseline
A584 pegIF-Na 4h
A584 difference

Appendix Figure S8

differentially expressed after pegIFNa, no difference between low ISG and high ISG patients
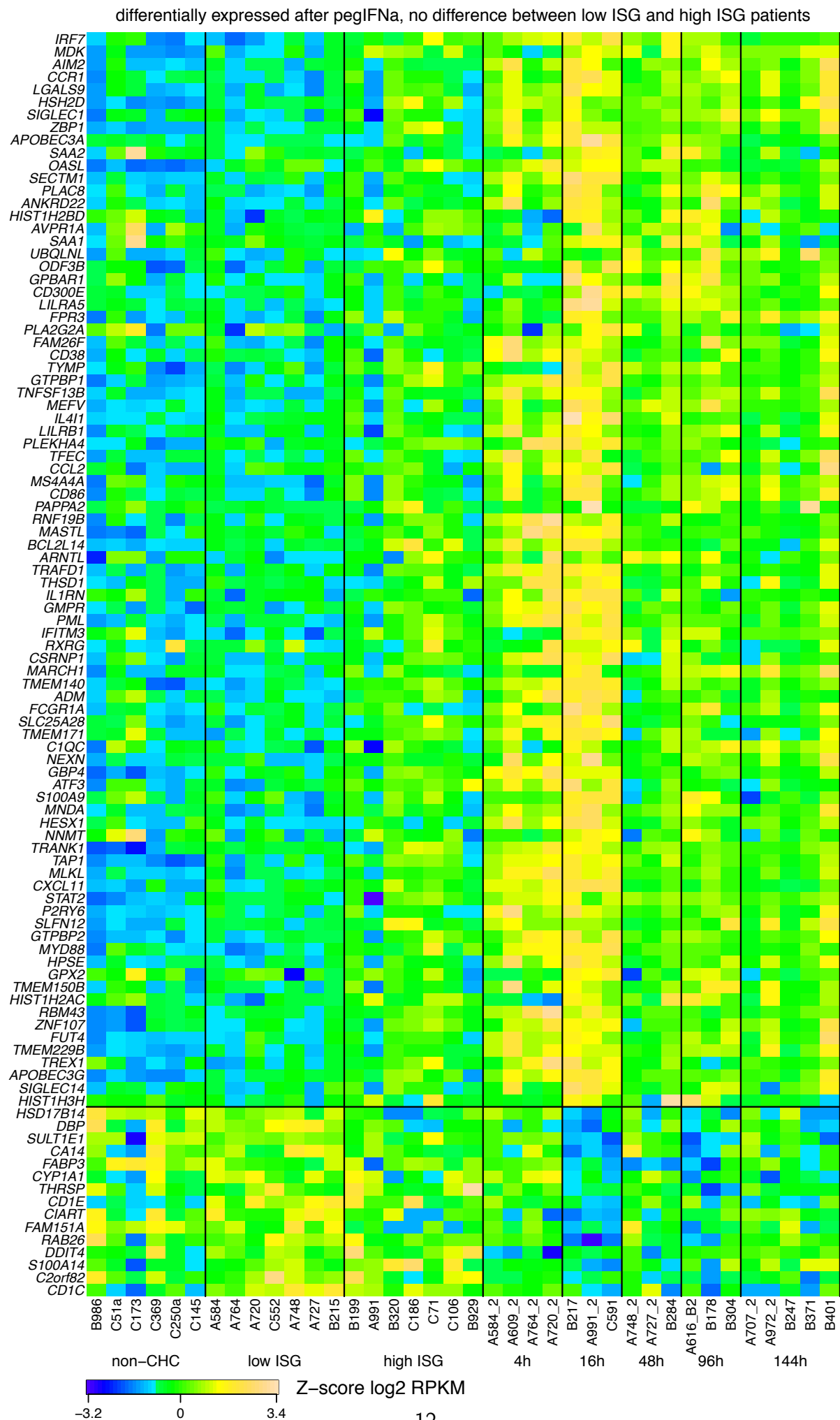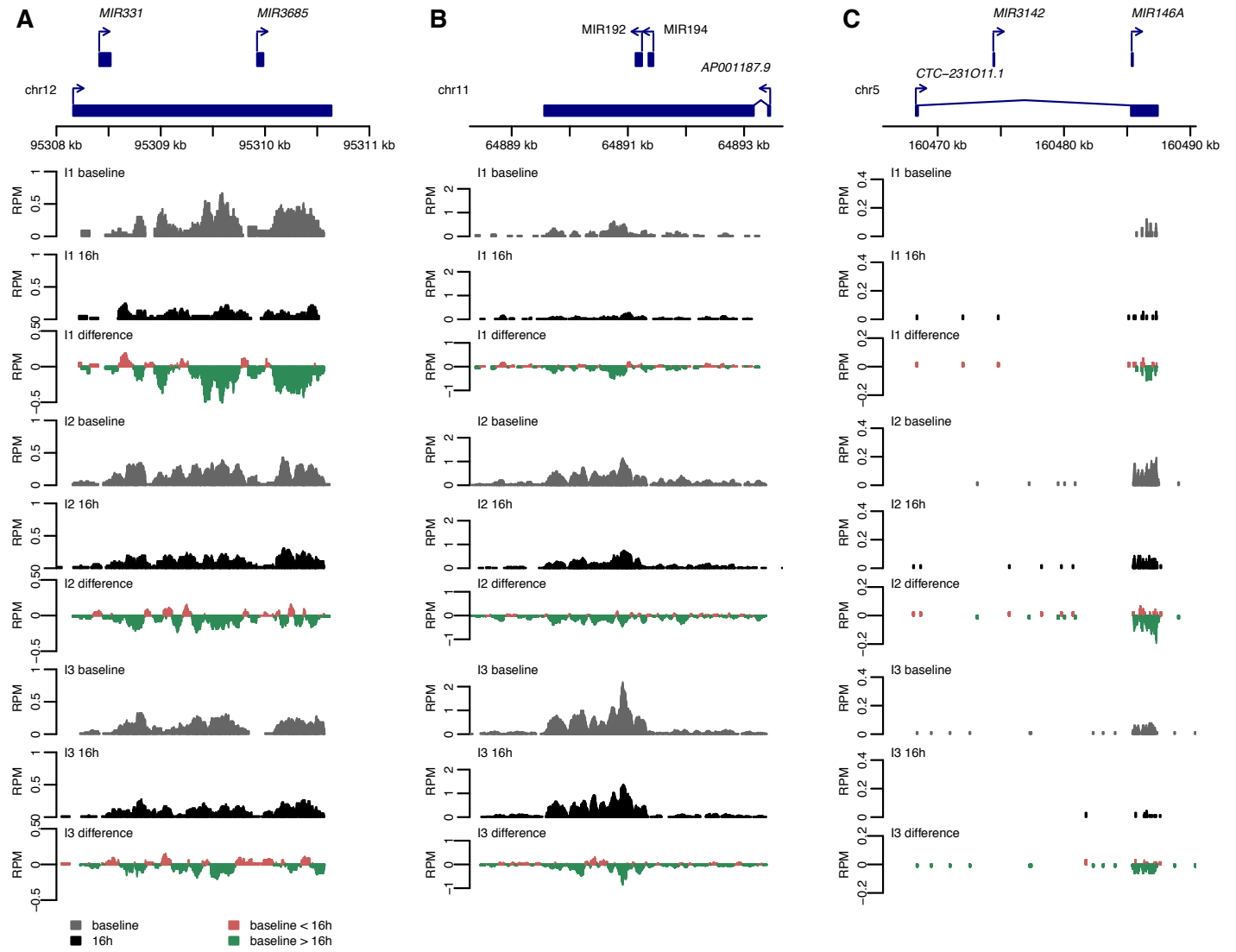
Z−score log2 RPKM

−3.2　0　3.4

Appendix Figure S9

Appendix Figure S11

Appendix Figure S12